ORIGINAL ARTICLE

Identification of high-risk and low-risk clusters and estimation of the relative risk of acute lymphoblastic leukemia in provinces of Iran during 2006–2014 period: A geo-epidemiological study

Sajjad Rahimi Pordanjani[1], Amir Kavousi[2], Babak Mirbagheri[3], Abbas Shahsavani[4,5], Koorosh Etemad[1]

[1]Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [2]Workplace Health Promotion Research Center, Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [3]Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran, [4]Environmental and Occupational Hazards Control Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [5]Department of Environmental Health Engineering, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran,

Background: The present study was conducted to determine the epidemiological status, identify high-risk and low-risk clusters, and estimate the relative risk (RR) of acute lymphoblastic leukemia (ALL) in provinces of Iran. **Materials and Methods:** This is an ecological study carried out using an Exploratory Multiple-Group design on 3769 children under 15 years of age with ALL from 2006 to 2014. Data analysis was performed using Mann–Whitney U, Global Moran's I and Kuldorff's purely spatial scan statistic tests at a significance level of 0.05. **Results:** The average annual incidence rate of ALL during 2006–2014 period was 2.25/100,000 children under 15 years of age. The most likely high-risk cluster with log-likelihood ratio (LLR) =327.47 is located in the southwestern part of Iran with a radius of 294.93 km and a centrality of 30.77 N and 50.83 E, which contained 1276 patients with a RR of 2.56. It includes Fars, Bushehr, Kohgiluyeh and Boyer-Ahmad, Khuzestan and Chahar Mahall and Bakhtiari provinces. On the other hand, the most likely low-risk cluster with 517 patients, and a RR 0.49 and LLR = 227.03 was identified in the northwestern part of Iran with a radius of 270.38 km and a centrality of 37.25 N and 49.49 E. It includes Zanjan, Qazvin, Gilan and East Azerbaijan, Ardabil, Alborz and Tehran provinces. **Conclusion:** High-risk clusters were observed in Southwestern, central, and eastern Iran, while low-risk clusters were identified in Northern and Western Iran.

Key words: Cluster, epidemiology, incidence, Iran, leukemia, risk

How to cite this article: Pordanjani SR, Kavousi A, Mirbagheri B, Shahsavani A, Etemad K. Identification of high-risk and low-risk clusters and estimation of the relative risk of acute lymphoblastic leukemia in provinces of Iran during 2006-2014 period: A geo-epidemiological study. J Res Med Sci 2021;26:18.

INTRODUCTION

Leukemia is a multifactorial disease with unknown etiology despite remarkable advances in the medical sciences.[1,2] According to the latest report issued by Iran's National Cancer Registration Program, which was recently released in 2019, the age standardized incidence rate of leukemia is estimated at 6.70/100,000 people.[3] High incidence and prevalence of leukemia lead to significant mortality and impose high diagnostic and therapeutic costs on Iran.[4-6]

Based on cell origin (lymphoid or myeloid) and duration of disease (acute or chronic), leukemia is divided into four main groups: Acute lymphoblastic leukemia (ALL), chronic lymphoid leukemia, acute myeloid

Access this article online

Quick Response Code:

Website: www.jmsjournal.net

DOI: 10.4103/jrms.JRMS_662_20

Address for correspondence: Dr. Koorosh Etemad, Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Velenjak St, Daneshjoo Blvd, Shahid Chamran Highway, Tehran, Iran. E-mail: etemadk@sbmu.ac.ir

Submitted: 30-Jun-2020 **Revised:** 08-Jan-2020 **Accepted:** 18-Sep-2020 **Published:** 27-Feb-2021

leukemia and chronic myeloid leukemia. The most common type of children cancer in the world and in Iran is acute lymphoblastic leukemia (ALL).[1,7] ALL accounts for about 80% of all childhood leukemia, imposing a significant public health burden on health care systems.[8,9] Children are the most valuable assets of any nation, whose health is a guarantee for future health of a country.

One of the earliest disease zones was proposed by John Snow in 1854, following the outbreak of cholera in Soho, London.[10] In epidemiology, a cluster refers to the accumulation of rare cases of disease or any health-related event such as cancer, congenital anomalies, suicide, and miscarriage which are interrelated in terms of time or space, and their number and accumulation is too large to be attributed to chance or coincidence. In clusters, disease cases are related to each other in terms of biological, pathological, environmental, social, or other specific conditions.[11]

Recent advances and high potentials of the geographic information system (GIS) and spatial epidemiology have opened a new window for epidemiologists, which allows planning for public health and new hypotheses.[11] Research on childhood leukemia can help identify environmental risk factors, providing new clues about the etiology of this disease.[12] On the other hand, investigating the spatial pattern of ALL occurrence can facilitate planning to promote disease surveillance system in high-risk areas, which promises to be an effective step in preventing the incidence and reducing the prevalence of ALL disease.

Therefore, considering the paucity of studies on the status of epidemiology and geographical pathology of ALL in Iran, the present research was conducted to not only determine the epidemiological status of the disease, but also estimate the relative risk (RR) and identifying high-risk and low-risk clusters of the disease. Overall, it demonstrates a more accurate assessment of the risk factors involved in the ALL cycle.

MATERIALS AND METHODS

Study design

This is an ecological study conducted using exploratory multiple-group design on all children under 15 years of age with ALL in Iran during 2006–2014 period. Given the exploratory nature of this study, it seeks to discover and describes disease patterns. Multiple-Group design also suggests that these studies are performed in several areas rather than one, and a comparison is drawn between the distributions of the disease in various locations.[13,14]

Target and study population

The target population of study comprised all children with ALL in Iran and the study population comprised children reported in Iran's National Cancer Registration Program with ALL in 2006–2014 years, who met the eligibility criteria of the study. The inclusion criteria were a known place of resident, diagnosis of ALL, and under 15 years of age.

Data resources

In this study, three main types of data sources were probe. The first source included data related to patients' demographic characteristics, which were obtained from Iran's National Cancer Registration Center for the period of 2006–2014. The second source consisted of data on the population at risk, or children under the age of 15 in the provinces of Iran, which was obtained from the Statistical Center of Iran. The third source contained data about patients' geographic coordinates, which were obtained from the latest information in the Google Map software and the georeferenced layers of Iran's provinces. Aggregate data were used in this study.

Spatial and descriptive characteristics of the study area:

Iran is a country in southwest Asia with 31 provinces and an area of 1,648,195 km^2 in the Middle East between 25°3′ and 39°47 north latitude and 44°5′ and 63°18′ east longitude. According to the last census in 2016, the population of Iran is estimated at 79927270 people [Figure 1].

Statistical analysis

Descriptive

In this section, first central tendency and dispersion indices of the disease were computed, and then the average annual incidence rate of the disease in Iran was estimated during 2006–2014. Mann–Whitney U test and SPSS software (ver. 18, International Business Machines Corporation (IBM), New York, USA) were used to compare the mean age of patients in male and female subjects. In the present study, statistical tests were performed at a significance level of 0.05.

Spatial autocorrelation and cluster identification

In order to evaluate the spatial distribution of disease incidence, first the cumulative incidence rate (CIR) of ALL was calculated for each province of Iran. In the next step, the Global Moran's I index was employed to assess the degree of spatial autocorrelation and the clustering tendency of ALL. The value of this index is between −1 and +1 with values close to +1 indicating a more clustered and values close to −1 denoting the dispersed distribution of the variable under study. Futhermore, values near zero suggests that the spatial distribution of the phenomenon under study is random.[15] Finally, Kuldorff's purely spatial scan statistic was used to identify the exact location of high-risk and low-risk clusters of ALL disease in Iran.

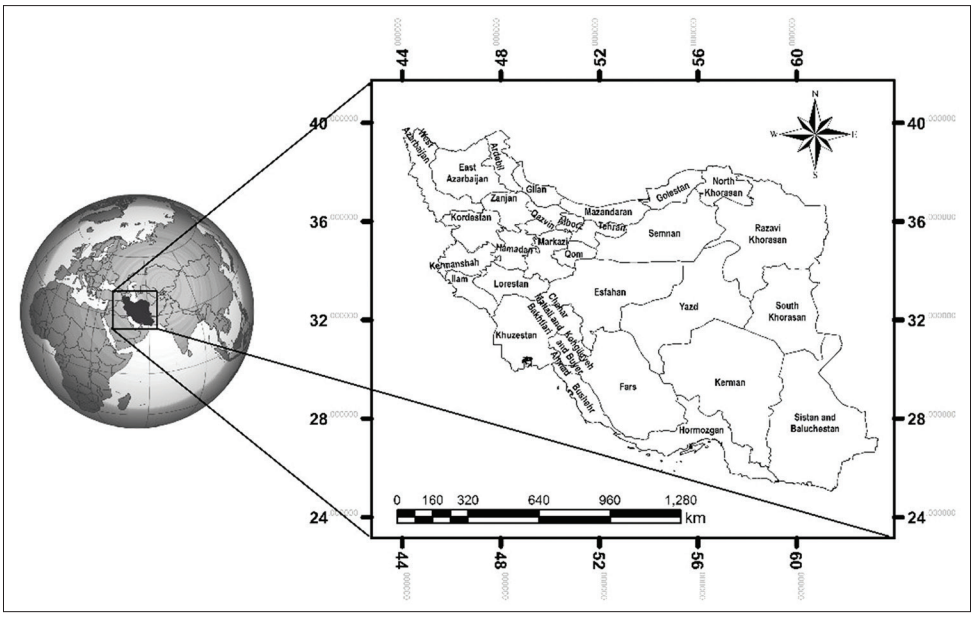

Figure 1: Geographical location of Iran

Kuldorff's spatial scan statistic

Due to its numerous advantages and applications, the Kuldorff's spatial scan statistic is widely used in the realm of public health and spatial epidemiology.[16] These advantages include the ability to analyze and identify retrospective and prospective clusters, and the ability to adjust confounders and covariates such as age and sex, to identify different types of Purely Spatial, Purely Temporal, and Spatio-Temporal clusters, to estimate the RR of disease in each of polygons and clusters, to detect and separate the most likely clusters and secondary clusters (lower probability), to detect high-risk and low-risk clusters simultaneously, to defined radius for the scanning window and to provide outputs in the format of programs such as ArcGIS, Google Earth, and Google Map.

This statistic is embedded in SaTScan software (ver. 9.6, Information Management Services, http://www.SaTScan.org, Boston, USA) that defines the scanning window in various circular, elliptical and irregular forms identifying clusters. In the present study, since the number of patients is a count and "discrete variable, in each possible location, the distribution discrete Poission and a circular window shape were used to identify clusters. In spatial exploration statistics, the null hypothesis is considered to be equal to the risk of disease inside and outside the window or cluster, and the alternative hypothesis is the high risk of disease inside the window or cluster. The risk of catching the disease in each window is calculated by log-likelihood ratio (LLR).[16] Under the Poisson assumption, the likelihood function for a window is as follows:

$$LLR = \left(\frac{c}{E[c]}\right)^{c}\left(\frac{C-c}{C-E[c]}\right)^{C-c}$$

Where C is the total number of cases, c is cases observed within the window, and E[c] is the expected number of cases inside the window under null hypothesis, which were adjusted for the covariates. Given that the analysis relies on the total number of observed cases, C-E[c] represents the expected number of cases outside the window.[16]

The window with maximized LLR is considered as the primary or most likely cluster (a cluster that is unlikely to be formed accidentally). Other windows with a lower LLR are considered as the secondary cluster (a cluster with a higher probability of being created accidentally).In this study, the report of secondary clusters is limited to nonoverlap secondary clusters.[16] Statistical significance was calculated by the Monte Carlo simulation method. The maximum number of Monte Carlo simulations was limited to 999.

As mentioned earlier, one of the advantages of this statistic is calculating the RR as one of the most important and widely used epidemiological indicators using the following method.

$$RR = \frac{c / E[c]}{(C-c)/(E[C]-E[c])}$$

Where c is the number of cases observed within window, E [c] is the number of cases expected within the window, C-c is the number of cases observed outside the window, E[C]-E[c] is the number of cases expected outside the cluster.[16] In the RR analysis, if the result of fraction is equal to one, the risk of disease within and outside window or cluster will be equal. Also, if the RR is >1, the risk of disease within the window will be higher than other areas outside the window, making that window a high-risk cluster. In addition, if RR is <1, the risk of

disease within the window will be lower than areas outside the window, making that window a low-risk cluster.[16,17]

Determining the optimal radius window for spatial cluster analysis

Before cluster analysis, it is necessary to define the radius of circular window, which is calculated from the number of at-risk populations in the study area at the user's discretion. Obviously, the size of clusters can never be encompassed by the maximum size defined for the window radius. By default, the hierarchical method of circular radius is considered to be at its maximum value, i.e., 50% of the population at risk. Then, in order to identify clusters, the circle radius is gradually increased from zero to a specified maximum value. In the default mode (hierarchal method), wide and large single clusters may be discovered, which are not useful from an epidemiological perspective.[16] To address this problem, there are two solutions: First, in the same hierarchical method, the user performs a sensitivity analysis of the percentage of population at risk and selects the optimum size of the circle radius according to the results of cluster identification. The second and more effective solution is to use the Gini optimized cluster collection method, which uses an accurate method to estimate the optimal radius size of the circle to identify clusters. It has recently been shown that the Gini coefficient can identify several nonoverlapping clusters in a more local, precise, and realistic way instead of identifying a large cluster.[16,18]

Therefore, to determine the size of circle radius and perform cluster analysis in the present study, first the default hierarchical method was used by setting up 50% of population at risk. Then, to identify clusters at a more local scale and to separate clusters that could be within clusters identified by the default hierarchal method, the Gini Optimized Cluster Collection method was utilized with an estimated optimal value of 31% of population at risk. Finally, the RR of ALL disease in all provinces of Iran was estimated and mapped in ArcGIS (ver. 10.8, Environmental Systems Research Institute, California, USA,). Figure 2 shows the flowchart of study stages.

RESULTS

Descriptive

In the present study, 3769 patients with ALL that met the inclusion criteria were entered into the study. The average annual incidence rate of ALL during the period of 2006–2014 was 2.25/100,000 children under 15 years of age. The mean and standard deviation of the age of patients were 5.90 ± 3.68, with a median age of 5 years and incidence peak of 2–5 years of age. As for gender, 1587 were female (42.1%) and 2182 were male (57.9%), indicating that the incidence of disease in males is 1.37 times higher than in females. The mean and standard deviation of age were 5.81 ± 3.61 in females and 5.97 ± 3.73 in males. The mean age was not significantly different between female and male patients ($P = 0.261$).

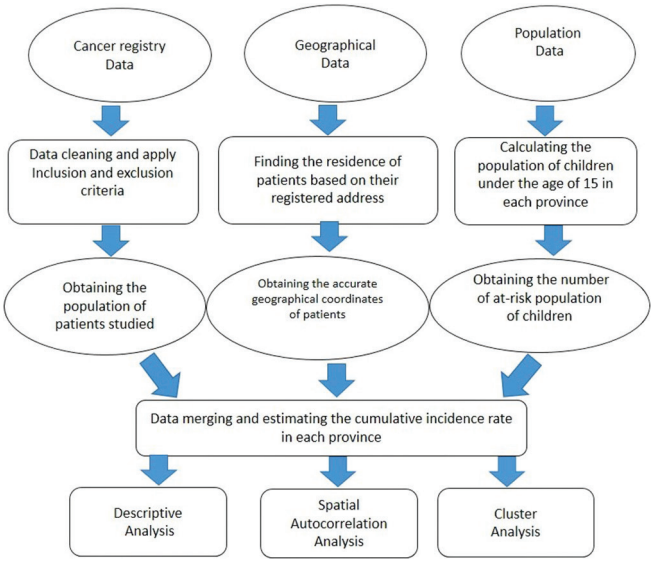

Figure 2: Flowchart of different stages of the study

Spatial autocorrelation

The value of the Global Moran's *I* was 0.358 and significant, which indicates the high intensity of autocorrelation and the high tendency of ALL for clustering in Iran, was significant at a confidence interval of 0.99 ($P = 0.0002$).

Identifying clusters using default hierarchical method

The results of cluster analysis using purely spatial scan statistic by considering 50% of population at risk for circular window size (hierarchical or default method) are shown in Figure 3. As can be seen, in this case, the spatial scan statistic identified only two clusters (a high-risk cluster and a low-risk cluster), which can be due to setting up the window radius at its highest level and limiting the report of secondary clusters in a nonoverlapping manner. High-risk clusters were identified in the southern and central provinces and low-risk clusters in the northern and western provinces of Iran. The complete epidemiological, statistical, demographical, and geographical characteristics of these clusters are shown in Table 1. The RR in the high-risk cluster was 2.7 and significant, suggesting that the risk of ALL in areas within cluster is 2.7 times higher than areas outside cluster (170% higher). Furthermore, the amount of RR in the low-risk cluster was 0.38 and significant, indicating that the risk of ALL in areas within the cluster is 0.38 times greater than areas outside the cluster (62% lower).

Identifying clusters by Gini optimized cluster collection

The results of Gini Optimized are depicted in Figure 4. As can be seen, the number of clusters detected by the Gini Optimized method is smaller than the figure identified by

Table 1: Characteristics of high-risk and low-risk acute lymphoblastic leukemia clusters shown in [Figure 3]

Cluster number in figure 3	Risk	Number of provincial involved in cluster	Coordinate center of cluster	Population size of cluster*	Radius (km) of cluster	Observed cases in cluster	Expected cases in cluster	Obs/Exp**	Relative risk	*P*	LLR***	Cluster type
1	High	13	27.02 N and 56.48 E	8,270,577	991.47	2644	1753.07	1.51	2.7	<0.001	430.27	Most likely
2	Low	16	35.68 N and 46,98 E	8,824,257	490.92	1033	1870.43	0.55	0.38	<0.001	386.42	Most likely

*Total population of children under the age of 15 in the cluster, **Number of cases observed divided by the number of cases expected in cluster. *** Log-Likelihood Ratio

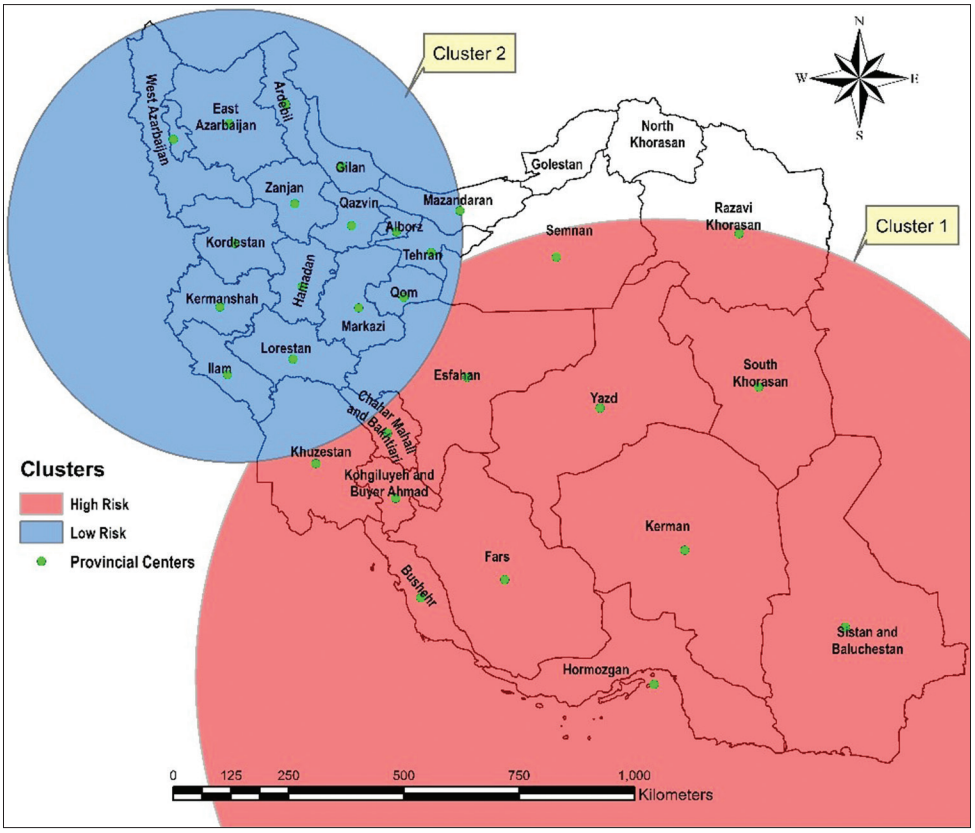

Figure 3: High-risk and low-risk ALL clusters in Iran using the hierarchical purely spatial scan statistic method

the hierarchical or default method, which is assumed to be due to the smaller size of the circular window radius (50% vs. 31%). A total of 5 important spatial clusters were discovered, including 2 high-risk clusters and 3 low-risk clusters.

The complete epidemiological, statistical, demographical, and geographical characteristics of the clusters discovered by Gini Optimized Cluster Collection method are shown in Table 2. According to the results of this table, it can be said that the most likely high-risk cluster with LLR = 327.47 is located in the southwestern part of Iran (center 30.77 N and 50.83 E) with a radius of 294.93 km. It includes Fars, Bushehr, Kohgiluyeh and Boyer-Ahmad Khuzestan and Chahar Mahall and Bakhtiari provinces [cluster 1 in Figure 4]. In this cluster, 1276 patients with ALL were identified and the RR of ALL was estimated (2.56). According to the results, the risk of ALL within this cluster is 2.56 times higher than areas outside the cluster. A high-risk secondary cluster was identified with LLR = 87.61 in the eastern Iran, which included the provinces of Yazd, Kerman, Razavi Khorasan, and South Khorasan.

On the other hand, the most likely low-risk cluster with LLR = 227.03 was discovered in the northwestern part of Iran (center 37.25 N and 49.49 E) with a radius of 270.38 km. It included the provinces of Zanjan, Qazvin, Gilan, East Azerbaijan, Ardabil, Alborz and Tehran [cluster 3 in Figure 4]. The number of patients identified in this cluster was 517 and the RR of ALL was estimated at 0.49, meaning that the risk of ALL within the cluster is lower than areas outside the cluster. Two secondary low-risk clusters were identified in the western and northeastern regions of Iran [clusters 4 and 5 in Figure 4] in which the RR was estimated at 0.3 and 0.62, respectively.

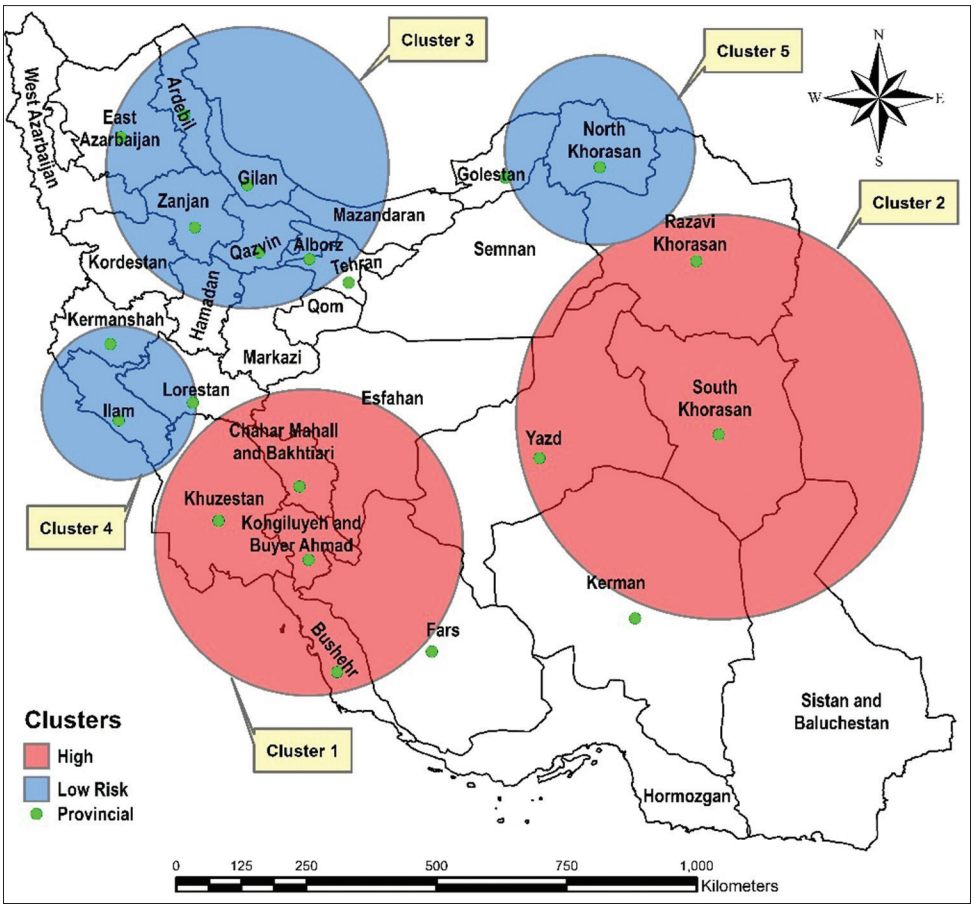

Figure 4: High-risk and low-risk clusters of ALL in Iran using the Gini Optimized Cluster Collection method

Estimation of the relative risk of all in all provinces of Iran

The RR of ALL incidence in all provinces was estimated and the results are shown in Table 3. These findings are consistent with the results of CIR in provinces, so that a high RR was observed in provinces with high CIR and a low RR in provinces with low CIR.

Four provinces with the highest RR were Fars (RR = 3.28), Kohgiluyeh and Boyer Ahmad (RR = 2.63), Yazd (RR = 2.23), and Khuzestan (RR = 1.83). On the contrary, four provinces with the lowest RR were Kermanshah (RR = 0.08), East Azerbaijan (CIR = 0.18), Zanjan (CIR = 0.22), and Ilam (CIR = 0.23). Finally, to illustrate the estimated RR values, the spot map of ALL was drawn in proportion to the RR at provincial scale [Figure 5].

DISCUSSION

The average annual incidence rate of ALL over a 9-year period was 2.25 per 100,000 children under 15 years of age. Comparing the results of this study with those reported in the literature showed that the incidence of disease in Iranian children is lower than in developed countries and similar to developing countries. The incidence of ALL tends to be higher in more developed countries with a higher human development index.[19] The incidence rate of ALL has been reported to be 6.4 in Mexico, 5.8 in Canada, 4.6 in the USA, and 4 in the UK per 100,000 children under the age of 15 years.[20] Other studies have estimated the incidence of the disease is between 2 and 7/100,000.[20-22] A 2017 population-based registry study of 62 countries found that leukemia in the United States, Europe, and Oceania is higher than Asia and Africa.[23]

However, it is predicted that by 2030, about 70% of new cases Leukemia will be reported in developing countries.[24] One reason for low incidence and mortality in sub-Saharan Africa is the lack of adequate diagnosis and treatment of patients.[25] The incidence and 5-year survival rate of ALL in whites is higher than the black people.[22] Down *et al*. documented an increase in ALL over a 25-year period in the United States.[22]

The results of age analysis of patients are consistent with other studies in this field. In most of reviewed studies, the highest age of disease incidence was reported at 2–5 years of age.[21,26] The incidence of ALL in children is about 4 times higher than in adults, and the disease has poor prognosis in adults. The highest incidence of ALL is shortly before school age.[22]

The high incidence of ALL in the early years of life may be due to poor immune systems of children and the body's abnormal response to common infections during this period.[27,28]

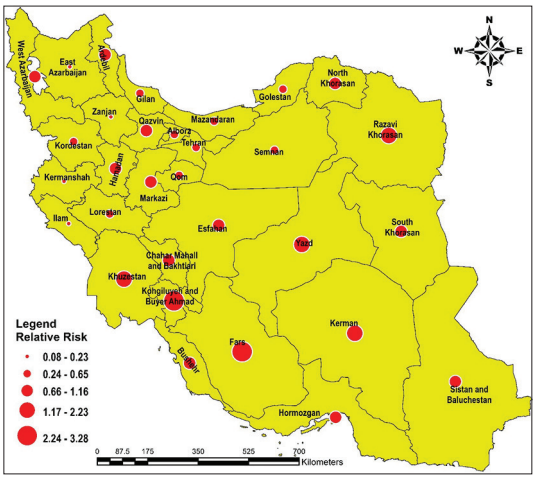

Figure 5: ALL spot map with respect to RR estimated in provinces

The sex ratio of male to female children was 1.37, but the average age of female and male patients was not statistically different. In the majority of studies, the incidence of the disease in boys has been reported to be higher than in girls.[1,20,21] According to a report published by the American Cancer Society in 2019, 55% of new cases and 56% of ALL-induced deaths were observed in boys.[29]

In short, in the study of demographic and epidemiological indicators of children with ALL in Iran, no unexpected and abnormal findings were achieved.

In exploring the spatial pattern of ALL incidence, two important points were identified: First, during global analyses, the disease had spatial autocorrelation and a

Table 2: Characteristics of high-risk and low-risk acute lymphoblastic leukemia clusters shown in Figure 4

Cluster number in figure 4	Risk	Number of provincial involved in cluster	Coordinate center of cluster	Population size of cluster*	Radius (km) of cluster	Observed cases in cluster	Expected cases in cluster	Obs/Exp**	Relative risk	*P*	LLR***	Cluster type
1	High	5	30.77 N and 50.83 E	2967044	294.93	1276	628.91	2.03	2.56	<0.001	327.47	Most likely
2	High	4	32.67 N and 59.22 E	2595277	386.5	856	550.11	1.56	1.72	<0.001	87.61	Secondary
3	Low	7	37.25 N and 49.49 E	5028019	270.38	517	1065.76	0.49	0.4	<0.001	227.03	Most likely
4	Low	3	33.12 N and 46.92 E	1027124	147.57	68	217.71	0.31	0.3	<0.001	73.69	Secondary
5	Low	2	37.40 N and 57.13 E	686464	181.31	92	145.51	0.63	0.62	<0.001	11.72	Secondary

*Total population of children under the age of 15 in the cluster, **Number of cases observed divided by the number of cases expected in cluster. *** Log-Likelihood Ratio

Table 3: Relative risk, observed and expected values of acute lymphoblastic leukemia in each of the provinces

Number	Province name	Obs cases*	Exp cases**	Obs/Exp***	Relative risk	No	Province name	Obs cases*	Exp cases**	Obs/Exp***	Relative risk
1	Alborz	56	91.71	0.61	0.60	17	Kordestan	34	81.38	0.42	0.41
2	Ardebil	68	68.38	0.99	0.99	18	Lorestan	53	95.82	0.55	0.55
3	Bushehr	92	79.53	1.16	1.16	19	Markazi	61	66.47	0.92	0.92
4	Chahar Mahall and Bakhtiari	37	49.49	0.75	0.75	20	Mazandaran	87	132.15	0.66	0.65
5	East Azarbaijan	33	176.25	0.19	0.18	21	North Khorasan	39	51.68	0.75	0.75
6	Esfahan	213	213.72	1.00	1.00	22	Qazvin	45	48.83	0.92	0.92
7	Fars	620	213.60	2.90	3.28	23	Qom	30	58.55	0.51	0.51
8	Gilan	54	109.21	0.49	0.49	24	Razavi Khorasan	459	318.98	1.44	1.50
9	Golestan	53	95.82	0.55	0.55	25	Semnan	15	28.28	0.53	0.53
10	Hamadan	93	86.99	1.07	1.07	26	Sistan and Baluchestan	176	186.53	0.94	0.94
11	Hormozgan	108	145.51	0.74	0.73	27	South Khorasan	45	48.83	0.92	0.92
12	Ilam	7	29.64	0.24	0.23	28	Tehran	249	517.93	0.48	0.44
13	Kerman	239	145.00	1.65	1.69	29	West Azarbaijan	143	161.41	0.89	0.88
14	Kermanshah	8	92.26	0.09	0.08	30	Yazd	113	51.47	2.20	2.23
15	Khuzestan	433	249.94	1.73	1.83	31	Zanjan	12	81.38	0.22	0.22
16	Kohgiluyeh and Buyer Ahmad	94	36.35	2.59	2.63						

*Observed cases in province, **Expected cases in province, ***Number of cases observed divided by the number of cases expected in province

high tendency for clustering. Second, during local analyses, high-risk clusters were identified in southwestern, central and eastern Iran and low-risk clusters in northern and western Iran.

The high autocorrelation and the emergence of spatial clusters in the incidence of ALL testify to the impact of environmental, geographical, infectious pathogenic and genetic susceptibility risk factors on the disease cycle. The results of this study, consistent with the literature on ALL clustering, suggested significant clusters for ALL occurrence. Researchers have reported environmental and infectious risk factors for the causes of ALL.[20,30,31]

A 2012 study by Stephen in Fallon, New York, found that leukemia cases were most common in the summer and underdeveloped areas and villages. They asserted that the spatio-temporal ALL distribution was abnormal and consistent with the involvement of infectious and viral agents.[32] In a 2013 study by Nyari *et al*. in Hungary, Chernobyl catastrophe was mentioned as a significant factor for ALL clustering.[33] Some viruses (HBV, HCV, HCG, HTLV-1) can be involved in leukemia due to their lymphothropism and acute and chronic infections in blood mononuclear cells leukocytes.[34-36] McNally *et al*.'s study in Manchester introduced infections as the cause of ALL clustering.[30] A 2017 study by Alvarez *et al*. found that ALL significantly increased clustering in Spain. Researchers cited children's environmental history, such as exposure to ionizing radiation and pesticides as the reason for this clustering.[1]

Of course, genetic susceptibility can also explain the difference in the incidence of ALL, with higher incidence reported in U. S. Hispanics and Latin Americans.[23] 3%–4% of leukemia cases are attributed to inherited genetic predisposition.[37]

According to the results of identifying high-risk and low-risk clusters, it can be stated that the incidence of the disease in low latitudes is much higher than high latitudes. It should be noted that the distribution of environmental risk factors varies in different latitudes and longitudes. In fact, the differences in the incidence of diseases in different places may be due to the environmental exposure in those areas.

One possible reason for high incidence of the disease in low latitudes and southern regions of Iran is greater exposure to ultraviolet radiation (UV) as a carcinogen in ALL, which is due to longer and direct sunlight. Another possible reason is the air pollution related to the oil-richness and the presence of vast gas resources in the southern regions of Iran, especially in Kohgiluyeh and Boyer-Ahmad and Khuzestan, Chahar Mahall and Bakhtiari provinces. In fact, in these areas, part of ALL incidence is due to the synergism interaction between the risk factors of geographical and environmental.

Patients' residence can be considered a valid indicator for assessing environmental and local exposures because children spend most of their times at home.[2] Geographical differences in ALL risk factors lead to changes in the spatial distribution and incidence of this disease.[2]

Chizhov *et al*. found that the association between solar activity and the risk of leukemia is important in raising awareness about the etiology of the cancer. In their 2018 study in Russia, they reported a significant linear relationship between leukemia and solar activity within 3 years after the birth of sick children (r = 0.567; P = 0.018).[38] In France. Astrid *et al*., in their 2015 study found a positive and significant association between the standardized incidence rate (SIR) of ALL and exposure to solar UV radiation. Researchers reported that UV contributes to the suppression of the immune system.[39] Elise *et al*., in a 2016 study on assessing evidence in the USA reported a high risk of leukemia in children with respect to unconventional oil and gas development. They posited that the health risks of UO and G development are novel and emerging in epidemiological studies. These findings call for further spatial analysis to detect leukemia using GIS.[40] Increased risk of ALL has also been reported in suburban highways,[41] near gas stations,[42] and adjacent to gasoline sources.[43]

CONCLUSION

The incidence of ALL in Iranian children is lower than in developed countries and similar to developing countries. ALL in boys is 1.37 times higher than in girls and is more common in the early years of life. ALL has a high spatial autocorrelation and has created cluster in Iran. High-risk clusters were observed in southwestern, central, and eastern Iran, and low-risk clusters were identified in northern and western Iran. In general, the condition of high-risk and low-risk clusters in Iran is very specific, indicating the synergism interaction between environmental, infectious, geographical, and genetic risk factors. Hence, it is suggested to strengthen disease surveillance systems to prevent the incidence and reduce prevalence, and to boost effective management and appropriate treatment of patients in high-risk areas of ALL in Iran.

Acknowledgments

This article is based on the results of PhD dissertation in Epidemiology (code of ethics: IR.SBMU.PHNS.REC.1398.143). We would like to thank the Deputy of Research and Technology of Shahid Beheshti University of Medical

Sciences, the Ministry of Health and Medical Education of Iran, the Statistical Center of Iran, Center for Remote Sensing and GIS Research of Shahid Beheshti University, whose assistance was crucial to the implementation of this study. Receiving support from the Center of Excellence in Analysis of Spatio-Temporal Correlated Data at Tarbiat Modares University is acknowledge.

Financial support and sponsorship

This work was supported by Shahid Beheshti University of Medical Sciences (SBUMS).

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Cárceles-Álvarez A, Ortega-García JA, López-Hernández FA, Orozco-Llamas M, Espinosa-López B, Tobarra-Sánchez E, *et al*. Spatial clustering of childhood leukaemia with the integration of the paediatric environmental history. Environ Res 2017;156:605-12.
2. Konstantinoudis G, Kreis C, Ammann RA, Niggli F, Kuehni CE, Spycher BD, *et al*. Spatial clustering of childhood leukaemia in Switzerland: A nationwide study. Int J Cancer 2017;141:1324-32.
3. Ministry of Health and Medical Education. Annual Report of Iranian National Population – Base Cancer Registry. Tehran: Mirmah; 2019.
4. Ehsani M, Shahgholi E, Hayati H, Kebriaeezadeh A, Nikfar S, Mehrvar A, *et al*. Cost pediatrics acute lymphoblastic leukemia based on ALL protocol. International Journal of Pediatrics, 2018;65:S113.
5. Keramatinia A, Mohseny M, Akbari ME, Mosavi-Jarrahi A, Monfared ED, Amanpour F, *et al*. Determinants of survival of common childhood cancers in Iran. J Res Med Sci 2018;23:101.
6. Zare N, Eskandari N, Mehrzad V, Javanmard SH. The expression level of hsa-miR-146a-5p in plasma-derived exosomes of patients with diffuse large B-cell lymphoma. J Res Med Sci 2019;24:10.
7. Reisi N, Azhir A, Hashemipour M, Raeissi P, Amini A, Moafi A. The metabolic syndrome in survivors of childhood acute lymphoblastic leukemia in Isfahan, Iran. J Res Med Sci 2009;14:111-6.
8. Figueroa SC, Kennedy CJ, Wesseling C, Wiemels JM, Morimoto L, Mora AM. Early immune stimulation and childhood acute lymphoblastic leukemia in Costa Rica: A comparison of statistical approaches. Environmental Research. 2020;182:109023.
9. Abudawood M. Diabetes and cancer: A comprehensive review. J Res Med Sci 2019;24:94.
10. Snow J. On the Mode of Communication of Cholera. London, England: John Churchill; 1855.
11. Kavousi A, Bashiri Y, Mehrabi Y, Etemad K, Teymourpour A. Identifying high-risk clusters of gastric cancer incidence in Iran, 2004-2009. Asian Pac J Cancer Prev 2014;15:10335-7.
12. McNally RJ, Alexander FE, Vincent TJ, Murphy MF. Spatial clustering of childhood cancer in Great Britain during the period 1969-1993. Int J Cancer 2009;124:932-6.
13. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008, p. 512-529.
14. Morgenstern H. Ecologic studies in epidemiology: Concepts, principles, and methods. Annu Rev Public Health 1995;16:61-81.
15. Cavalcante ACP, de Olinda RA, Gomes A, Traxler J, Smith M, Santos S. Spatial modelling of the infestation indices of *Aedes aegypti*: An innovative strategy for vector control actions in developing countries. Parasit Vectors 2020;13:197.
16. Kulldorff M. SaTScanTM User Guide for Version 9.6, 2018; 2018. (https://www.satscan.org/).
17. Hohl A, Delmelle EM, Desjardins MR, Lan Y. Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. Spat Spatiotemporal Epidemiol 2020;34:100354.
18. Kim J, Jung I. Evaluation of the gini coefficient in spatial scan statistics for detecting irregularly shaped clusters. PLoS One 2017;12:e0170736.
19. Hubbard AK, Spector LG, Fortuna G, Marcotte EL, Poynter JN. Trends in international incidence of pediatric cancers in children under 5 years of age: 1988-2012. JNCI Cancer Spectr 2019;3:pkz007.
20. Tlacuilo-Parra A, Garibaldi-Covarrubias R, Romo-Rubio H, Soto-Sumuano L, Ruiz-Chávez CF, Suárez-Arredondo M, *et al*. Geographical distribution and cluster detection of childhood leukemia in the metropolitan area of Guadalajara, Mexico. Rev Invest Clin 2017;69:159-65.
21. Bahoush G, Nojoomi M. Frequency of cytogenetic findings and its effect on the outcome of pediatric acute lymphoblastic leukemia. Med Arch 2019;73:311-5.
22. McNeil DE, Coté TR, Clegg L, Mauer A. SEER update of incidence and trends in pediatric malignancies: Acute lymphoblastic leukemia. Med Pediatr Oncol 2002;39:554-7.
23. Steliarova-Foucher E, Colombet M, Ries LAG, Moreno F, Dolya A, Bray F, *et al*. International incidence of childhood cancer, 2001-10: A population-based registry study. Lancet Oncol 2017;18:719-31.
24. Pérez-Cuevas R, Doubova SV, Zapata-Tarres M, Flores-Hernández S, Frazier L, Rodríguez-Galindo C, *et al*. Scaling up cancer care for children without medical insurance in developing countries: The case of Mexico. Pediatr Blood Cancer 2013;60:196-203.
25. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, *et al*. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-86.
26. Hein D, Borkhardt A, Fischer U. Insights into the prenatal origin of childhood acute lymphoblastic leukemia. Cancer Metastasis Rev 2020;39:161-71.
27. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer 2006;6:193-203.
28. Ross JA, Severson RK, Swensen AR, Pollock BH, Gurney JG, Robison LL. Seasonal variations in the diagnosis of childhood cancer in the United States. Br J Cancer 1999;81:549-53.
29. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7-34.
30. McNally RJ, Alexander FE, Birch JM. Space-time clustering analyses of childhood acute lymphoblastic leukaemia by immunophenotype. Br J Cancer 2002;87:513-5.
31. Agost L. Analysis of spatial-temporal clusters of childhood cancer incidence in the province of Córdoba, Argentina (2004-2013). Arch Argent Pediatr 2016;114:534-43.
32. Francis SS, Selvin S, Yang W, Buffler PA, Wiemels JL. Unusual space-time patterning of the Fallon, Nevada leukemia cluster: Evidence of an infectious etiology. Chem Biol Interact 2012;196:102-9.
33. Nyari T, Ottoffy G, Bartyik K, Thurzó L, Solymosi N, Cserni G, *et al*. Spatial clustering of childhood acute lymphoblastic leukaemia in Hungary. Pathol Oncol Res 2013;19:297-302.
34. Takai S, Tsurumi H, Ando K, Kasahara S, Sawada M, Yamada T, *et al*. Prevalence of hepatitis B and C virus infection in haematological malignancies and liver injury following chemotherapy. European Journal of Haematology2005;74:158.
35. Bagheri K, Yaghobi R, Karimi MH, Mirzaei M, Ramzi MJ. Determination of serologic and molecular prevalence of hepatitis Type B, C, and G infections in patients with hematological

malignancy in the South of Fars province, Iran. J Fasa Univ Med Sci 2011;1:168.
36. Timonen TT. A hypothesis concerning deficiency of sunlight, cold temperature, and influenza epidemics associated with the onset of acute lymphoblastic leukemia in northern Finland. Ann Hematol 1999;78:408-14.
37. Bloom M, Maciaszek JL, Clark ME, Pui CH, Nichols KE. Recent advances in genetic predisposition to pediatric acute lymphoblastic leukemia. Expert Rev Hematol 2020;13:55-70.
38. Chizhov AY, Pinaev S. Effects of solar radiation and woodsmoke on risk of childhood leukaemia: System analysis. Radiation Risk 2018;27:87-94.
39. Coste A, Goujon S, Boniol M, Marquant F, Faure L, Doré JF, *et al*. Residential exposure to solar ultraviolet radiation and incidence of childhood hematological malignancies in France. Cancer Causes Control 2015;26:1339-49.
40. Elliott EG, Trinh P, Ma X, Leaderer BP, Ward MH, Deziel NC. Unconventional oil and gas development and risk of childhood leukemia: Assessing the evidence. Sci Total Environ 2017;576:138-47.
41. Spycher BD, Feller M, Röösli M, Ammann RA, Diezi M, Egger M, *et al*. Childhood cancer and residential exposure to highways: A nationwide cohort study. Eur J Epidemiol 2015;30:1263-75.
42. Brosselin P, Rudant J, Orsi L, Leverger G, Baruchel A, Bertrand Y, *et al*. Acute childhood leukaemia and residence next to petrol stations and automotive repair garages: The ESCALE study (SFCE). Occup Environ Med 2009;66:598-606.
43. McKenzie LM, Allshouse WB, Byers TE, Bedrick EJ, Serdar B, Adgate JL. Childhood hematologic cancer and residential proximity to oil and gas development. PLoS One 2017;12:e0170423.